# EAHR 811 Jim Thompson SQR3 Chapters 3 and 4

## Chapter 3 – Alternative Views of Evaluation

### Diverse Conceptions of Program Evaluation

* Early days evaluation troubled by definition and ideological disputes.
  + If based on professional judgment 🡪 expert assessment.
  + If comparing performance to objectives 🡪 standardized constructs.
  + If a decision oriented approach 🡪 judge which decisions were best.
  + Analyze & interpret data 🡪 evaluate worth of curriculum for recommendation.

### Origins of Alternative Views of Evaluation

* “The diversity of evaluation approaches...has arisen from the varied backgrounds and worldviews of their authors (Fitzpatrick et al, 2004.)” Orientations, methods, practical preferences.
* Baker and Naomi (1996) proposed (1) experimentation, (2) measurement, (3) system analysis, and (4) interpretative approaches.
* Objectivist & Subjective Epistemology
  + Objectivism 🡪 requires that evaluation information be ‘scientifically objective.’ Evaluation procedures are externalized.
  + Subjectivism 🡪 “an appeal to experience” rather than scientific method. Internalized by the evaluator in ways that may not be reproducible by others.
  + Both methods have their advantages and disadvantages based on their root premise. The debate surrounding logical positivism is a demonstration of this root debate.
* Utilitarian vs Intuitionist-Pluralist Evaluation
  + Utilitarian 🡪 determine value by assessing the overall impact of a program.
  + Intuitionist-Pluralist 🡪 program value depends on impact program has individual.
* Impact of Philosophical Differences 🡪 acceptable program evaluation depends on the epistemology philosophy.
* Methodological Backgrounds & Preferences 🡪 quantitative and qualitative data.
  + Qualitative Data 🡪 non-numerical (a simplistic view) narration, essay
  + Quantitative Data 🡪 numerical, statistics to summarize data.
* Disciplinary boundaries & evaluation boundaries 🡪 “Evaluation is not a traditional discipline but a trans-discipline that ....cuts across disciplines...evaluators must avoid the luxury of remaining within any one discipline.
* Different Metaphors of Evaluation 🡪 account for much of the variation in evaluation approaches.
* Responding to different needs 🡪 comprehend the wide range of needs, identify what is useful in each approach, avoid irrelevant approaches.
* Practical considerations 🡪 value judgment? Political roles! Evaluator’s prior experience, expertise, variety of evaluation approaches used (simple ↔ complex)

### A Classification Schema for Evaluation Approaches (5 Categories)

1. Objective-oriented approach 🡪 specify goals & objectives and determine the extend to which they have been met.
2. Management –oriented approaches 🡪 identifying and meeting the informational needs of managers (decision making.)
3. Consumer-oriented approaches 🡪 developing evaluative information on ‘products.’
4. Expertise-oriented approaches 🡪 direct application of professional expertise.
5. Participant-oriented approaches 🡪 involvement of participants central in determining the values, criteria, needs, data, and conclusions for the evaluation.

## Chapter 4 – Objective Oriented Evaluation Approaches

“The information gained from an objective-oriented evaluation could be used to reformulate the purposes of activity, the activity itself, or the assessment procedures and devices used to determine the achievement of purposes (Fitzpatrick, 2004.)”

### Developers of Objective Oriented Evaluation Approach/Contributions (Tyler)

* Tyler’s approach to evaluation followed these steps: (logical, systematic, scientific, etc)
  1. Establish broad goals or objectives.
  2. Classify the goals or objectives.
  3. Define objectives in behavioural terms.
  4. Find situations in which achievement of objectives can be shown.
  5. Develop or select measurement techniques.
  6. Collect performance data.
  7. Compare performance data with behaviourally stated objectives.
* Tyler’s six categories for American schools; Acquisition of information, Development of work habits and study skills, Development of effective ways of thinking, Internalization of social attitudes, interests, appreciations, and sensitivities; Maintenance of physical health, Development of a philosophy of life.
* Handbook of Educational variables has modified the purposes; intellectual, emotional, physical and recreational, aesthetic and cultural, moral, vocational, and social.
* Saunders & Cunningham (1974) ...Logical Methods:
  1. Need for accomplishing the goal or objective is critical consideration.
  2. Examine the consequences of accomplishing the goal or objective.
  3. If a goal or objective conflicts with higher-order values, it may create more problems.
* Empirical methods;
  1. Collecting group data to describe judgments about the value of a goal.
  2. Arranging for experts, hearings, or panels to review and evaluate potential goals or objectives.
  3. Conducting content studies of archival records.
  4. Conducting a pilot study to see whether the goal is attainable.
* Metfessel & Michael’s Evaluation Paradigm 🡪 Tyler heavily influenced their eight step evaluation process.
* Provus’s Discrepancy Evaluation Model 🡪 continuous information-management process.
* Evaluation Cube 🡪 3D framework for analyzing the objects of community based programs.
* Logic Models 🡪 long term vision of how participants will be better off (ultimate outcome.)
* Program Theory 🡪 delivery of program and appearance of outcomes of interest.

### How has the Objective Oriented Evaluation been used.

* Dominated evaluation since the 1930’s,
* Standards based, objective referenced testing movements (1970’s.)

### What are the strengths and limitations of the Objective Oriented Evaluation Approach

* Easy to understand, follow, and implement 🡪 produces relevant information to the program.
* Goal-free Evaluation 🡪 the evaluator purposely avoids becoming aware of program goals.
  + Predetermined goals do not permit narrowing the focus of the study.

# Personal Reflection

I found these two chapters interesting as they relate to a course I just completed two weeks ago in Vancouver. The course was designed to meet the criteria as an Approved Check Pilot. There were 24 participants in this five day course, nine of us were ‘initial ACP candidates.’

On the first day of the course the goal and mission of an ACP was instilled in us nine initial candidates. Be **objective**, measure training outcomes by evaluating pilots during a pilot proficiency ride, a ‘snapshot’ of pilot competency conducted once every year or six months depending on aviation activity.

This mantra was what I expected, as a pilot being tested during a Pilot Proficiency Check (PPC) and as an instructor I am acutely aware of the good, bad, and ugly associated with an objectively based ‘snapshot’ evaluation. We initial candidates soaked up this philosophy absorbing every word of the objective, repeatable, demonstrative, external, philosophy.

As the course continued throughout the week, we were an isolated group of 24 ‘experienced’ pilots who would be conducting check rides on fellow pilots to ensure the safety of the operation and promote confidence in pilots. Some of the exercises we did were evaluation exercises on structured check ride scenarios. Transport Canada had provided the class with videos and written synopsis of actual ride occurrences.

The facilitator had the class review the written synopsis and videos, and evaluate the crew performance using a scale of 1 to 4 where one is a failure, and four is god. To the surprise of several people (both initial and recurrent) a consensus of the crew performance could not be attained by the class. While most of the class believed the performance was a 1(fail) or 2 (good enough, but we need to talk) evaluation the reasons varied.

This was the opportunity for the facilitator to emphasize that while the evaluation is objective your interpretation of the objectives is based on your expertise, a subtle way to introduce the group to **varied approaches** to evaluation. In other words use objective criterion to guide your subjective evaluation. This was something new, a definite change in how to evaluate performance. In the past subjective evaluations have been used, but never spoken of. This was the first time the discussion of subjective assessment using objective criteria had ever been mentioned in such an aviation forum, it had long been practised but never spoken!

The course was interesting because of the discussion concerning evaluating people skills and abilities as part of the ‘unofficial’ performance criteria. There are several examples in check rides where the performance objective was met, but it was not pretty. Again, this was the first time where a candid discussion of how to evaluate flightcrew situational awareness, cockpit resource management, and standard operating procedures had been discussed in an open forum. Currently none of these capabilities can be specifically evaluated. The facilitator and recurrent ACP participants acknowledged that this was a function of a check pilot to find an objective to which they could associate undesirable performance. For example, an undesirable performance, obviously due to a lack of communication between flightcrew, that results in failure of the objective. While the Check Pilot realizes that the lack of communication was a key factor in the poor performance they would have to fail the crew on an objective criteria.

A significant amount of time in the course discussed the proper use of terminology to describe performance. For legal reasons, less is more, and terminology was crucial when considering who would be reading your pilot evaluation following a fatal accident. If a candidate did fail an objective then the facilitator recommended a very explicit detail of how the candidate failed. For example “Failed to maintain Minimum Descent Altitude by descending 200 feet below target MDA...” This was opposed to a subjective evaluation of the same performance where the Check Pilot might consider passing the pilot and provide a written comment “momentary deviation below MDA, promptly recovered.” Given the personalities attracted to the Check Pilot world, this discussion was not received in a gracious forum. The fact that someone would pass or fail this same individual created an active debate which was enjoyable.